أخلاقيات الذكاء الاصطناعي

يسري مصطفى

الأخلاق والذكاء صفتان أساسيتان للإنسانية، وهما موجودتان منذ قديم الزمان؛ فالأخلاق تعمل كإطارٍ للقِيَم والتنظيم الاجتماعي، وتوجِّه السلوك البشري والتفاعلات، في حين يتميز الذكاء بالفهم والمعالجة والإبداع والقدرة على حل المشكلات. والتفاعل بين الذكاء والأخلاق مُعقَّد ومُتعدِّد الأوجه؛ فهي ليست مُجرَّد علاقةٍ خَطِّيَّةٍ بل تفاعلية ونسبية تتطور، وتتغير بتغيُّر أحوال البشر وبيئاتهم الاجتماعية والسياسية والقِيَمِيَّة. ومن ناحيته، يلعب العقل، كرَحِمٍ للتفكير والمعرفة، دورًا حاسمًا في توليد القِيَم والمبادئ الأخلاقية. وهذا الإطار الأخلاقيّ الجماعيّ يرشد –بدوره– الذكاء الفرديَّ وقد ينمّيه أو يحدُّه. وعلى مر التاريخ، أنتج العقل مجموعةً واسعةً من المعرفة والتقدُّم العلميّ، وبموازاة ذلك أنتج مبادئ أخلاقيةً لترسيم الحدود المنظِّمة لإنتاج هذه المعرفة وتطبيقها. وكانت هذه العلاقة الديناميكية واضحةً بشكلٍ خاصّ في جميع مجالات الحياة تقريبًا، وسار تاريخ البشر في مساراتٍ متعدِّدةٍ بين العقل والأخلاق؛ فثمة توافقات وصراعات بين سلطة العقل (أو العلم) وسلطة الأخلاق (أو القِيَم). وبدخول أزمنة الحداثة اتسعت تطبيقات العلوم والتكنولوجيا ومعها زادت المنافع والمخاطر وتعاظمت سلطة العقل والعلم، فزاد الاهتمام بالأخلاقيات كحصون قِيَمِية وقائية، أو كأساسٍ للتشريعات والسياسات ومواثيق الشروف ومدوَّنات السلوك.

ومع التطورات التكنولوجية والعلمية التي اكتسحت عالمنا منذ منتصف القرن الماضي، بات الأمر وكأن هناك ملاحقة أخلاقية لتطورات متسارعة تشق طريقها بلا هوادة فتفرز مخاطر آنية ومستقبلية. على سبيل المثال، دفع تطوير التكنولوجيا النووية والهندسة الوراثية وتكنولوجيا المعلومات المجتمع إلى التعامل مع أسئلة أخلاقية عميقة تتعلق بالسلامة والخصوصية والآثار الأخلاقية للتدخل البشري في العمليات الطبيعية. وفي السنوات الأخيرة، أدخل ظهور الذكاء الاصطناعي بُعدًا جديدًا للعلاقة بين الذكاء والأخلاق. فلأول مرة في التاريخ، يكون هناك ذكاء موازٍ للذكاء البشري، وأكثر من ذلك، ذكاء اصطناعي بإمكانه تجاوز الذكاء البشري في مجالات مختلفة. ويثير هذا التقدم غير المسبوق أسئلةً أخلاقيَّةً حاسمةً حول طبيعة الذكاء نفسه والمخاطر المرتبطة بإنشاء كيانات قد تعمل بشكلٍ مستقلٍّ عن الأُطُر الأخلاقية البشرية. ومع تزايد تعقيد أنظمة الذكاء الاصطناعي، فإن إمكانية عملها بطرق لا تتماشى مع القِيَم الإنسانية باتت مصدر قلق. وإلى جانت المخاوف الأخلاقية الآنية تتعاظم الحاجة إلى إجراء تقييمات أخلاقية استباقية لما يمكن أن تصل إليه تقنيات الذكاء

الاصطناعي في المستقبل القريب. ويتعين على هذه التقييمات أن تعالج أسئلة مثل: كيف نضمن أن أنظمة الذكاء الاصطناعي تتوافق مع القيم الإنسانية؟ وما هي الضمانات التي يمكن تنفيذها لمنع سوء الاستخدام أو العواقب غير المقصودة؟ وكيف نتعامل مع الآثار الأخلاقية المترتبة على تفويض اتخاذ القرار إلى الآلات؟ إن إثارة هذه الاعتبارات الأخلاقية يؤكد على أهمية موضوع "أخلاقيات الذكاء الاصطناعي". ومع الاستمرار في الابتكار ودمج الذكاء الاصطناعي في جوانب مختلفة من الحياة تزداد نبرة الانبهار ومنسوب المخاوف. وإذا كانت المخاوف مُبرَّرةً، فهل يمكن أن توفر الأخلاق لنا سُبُل الخلاص من المخاطر الواقعيَّة والمُحتَمَلة؟

في البداية، من الأهمية بمكان أن ندرك أن الخصائص المتأصِّلة لتكنولوجيا الذكاء الاصطناعي تضيف طبقة جديدة من التعقيد إلى المناقشات الجارية حول التفاعل بين الأخلاق والعلم والتقدُّم التكنولوجيّ في مجالات الحوسبة والاتصالات. فعلى النقيض من أي ابتكاراتٍ تكنولوجيةٍ سابقةٍ، يطرح الذكاء الاصطناعيُّ تحدياتٍ فريدةً تتحدَّى الأطر التنظيميَّة التقليديَّة. ويرجع هذا في المقام الأول إلى أن أنظمة الذكاء الاصطناعي ليست مجرد أدوات؛ فهي قادرة على التعلُّم والتكيُّف والتطور بطرق تعقِّد فهمنا لتداعياتها وطرق حوكمتها. على سبيل المثال، يثير التقدم السريع لتقنيات الذكاء الاصطناعي تساؤلات حول المساءلة والشفافية. من المسؤول عندما يتخذ نظام الذكاء الاصطناعي قرارًا يؤدي إلى ضرر؟ كيف نضمن أن تعمل هذه الأنظمة بشكل عادل ومن دون تحيز؟ وهذا وضع قائم بالفعل، ولكننا نعلم أن هناك من يحتكر ومن يوظف ومن يستفيد، ولكن الخوف من المستقبل قد يكون أعظم من الخوف المتحقق الآن، بسبب التنبؤات المتعلقة بقدرة أنظمة الذكاء الاصطناعي أن تعمل بشكل مستقل في المستقبل لقريب، أي أن تصبح خارج السيطرة. إنها فكرة غائمة ولكنها مُحتَمَلة. يخلق هذا الافتقار إلى الوضوح حاجة ملحة لأطر أخلاقية جديدة ونهج تنظيمي قادر على مواكبة تطور التكنولوجيا.

وعلاوة على ذلك، فإن تأثيرات الذكاء الاصطناعي تُعقِّد حل هذه القضايا الأخلاقية والسياسية لكونها سريعة التطور والانتشار مما يؤثر على مختلف قطاعات المجتمع بطرق عميقة. من الألعاب والترفيه إلى الموسيقى والمراقبة والحرب والبحث العلمي والرعاية الصحية والإدارة البيئية والتعليم والحوكمة وأنظمة العدالة، وهكذا فإن الذكاء الاصطناعي يدخل في نسيج الحياة اليومية بشكلٍ متزايدٍ. كلٌّ من هذه المجالات تقدِّم معضلاتٍ وتحدِّياتٍ أخلاقيَّةً فريدةً تتطلب دراسةً متأنيةً واستجاباتٍ مخصَّصةً. وما يزيد من المخاوف، هو السؤال عمَّن يسيطر على هذا التطور الهائل؟ فثمة احتكارات وثمة أغلبية مُهمَّشة، وسوف تزداد تهميشًا بسبب سرعة الابتكارات والتحولات. وغالبًا ما يتميز هذا التطور السريع بسباقٍ تنافسيٍّ بين القوى العالمية الكبرى، وخاصةً الولايات

المتحدة والصين وأوروبا. هذه المنافسة، التي يشار إليها غالبًا باسم "سباق التسلح بالذكاء الاصطناعي". ومع سعي الدول إلى التفوق على بعضها البعض في قدرات الذكاء الاصطناعي، فإن احتمالات إساءة الاستخدام والعواقب غير المقصودة تتزايد، مما يزيد من المخاطر على الأمن العالمي والحوكمة الأخلاقية. ومع التقدم إلى الأمام، سيكون التحدي هو تحقيق التوازن بين الإمكانات الهائلة للذكاء الاصطناعي والحاجة إلى الرقابة والمساءلة، وضمان أن يعود تطوره بالنفع على المجتمع ككل. وهذا الهدف موضع شكٍّ، لسبب بسيط وهو أن مسار تطور الذكاء الاصطناعي لا يعطي مؤشرات واضحة على إمكانية حوكمته بشكل عادل ورشيد.

وبالعودة إلى سؤال الأخلاق، فإن الأمر يبدو وكأن مصطلح الأخلاق واضحٌ وشفافٌ ومحلُّ اتفاقٍ، ولكن في حقيقة الأمر فإن فكرة الأخلاق في حدِّ ذاتها خلافيَّةٌ وإشكاليَّةٌ وتثير تساؤلاتٍ من منظور طبيعتها وتجانسها. كما أن علاقة الأخلاق بالذكاء الاصطناعي تثير تساؤلات عملية من منظور الفاعليَّة والملاءَمَة.

فمن ناحية أولى، يتَّسم المشهد الأخلاقي بغياب ملحوظ للإجماع العالمي؛ حيث لا تزال هناك اختلافات كبيرة في القِيَم بين الثقافات فيما يتعلق بما نشير إليه بالقِيَم الأخلاقيَّة. ويثير هذا الاختلاف أسئلةً حاسمةً حول أسس المبادئ الأخلاقية عبر المجتمعات المختلفة. على سبيل المثال، هل يمكننا تحديد اتفاقٍ أخلاقيٍّ مُشتَرَك بشأن قضايا أساسية مثل حقوق المرأة، أو حرية الفكر والمعتقد، أو الحق في الخصوصية؟ في الحقيقة، إن هذه التوافقات غير موجودة. وفي حين توفِّر لغة الحقوق إطارًا يحدِّد ما يُعتبر مقبولًا أو إلزاميًّا في سياقاتٍ مختلفةٍ، فمن الضروريّ أن ندرك أن الحقوق في حدِّ ذاتها لا تعادل الأخلاق، كما أنها تمثل مجموعةً من المبادئ والمعايير التي قد لا تتوافق مع القيم المتنوعة والأنظمة الأخلاقيَّة الموجودة عالميًّا. ومع تعمُّقنا في عالم التكنولوجيا، تصبح مناقشة الاعتبارات الأخلاقية أكثر تعقيدًا؛ فمعضلة الأخلاق لا تتعلق فقط بمدخلات الذكاء الاصطناعي، بل بكامل المنظومة وما تتضمنه من عمليات ومُخرَجَات. قد يكون من السهل نسبيًّا معالجة قضايا مثل تجنب التحيز، أو الامتناع عن الأذى، أو احترام حقوق الملكية الفكرية في السياقات التكنولوجية. ولكن التعقيد يتصاعد بشكلٍ كبيرٍ عندما نفكر في الآلات التي تُوصَف بالذكية، الآلات التي يُعتَمَد عليها بشكلٍ متزايدٍ في اتخاذ القرارات الحاسمة، وإنشاء الصور، والأنشطة البحثية، فهي تحتاج إلى مُدخَلاتٍ أخلاقيَّةٍ، ولكن هذه المدخلات ذاتها ليست محلَّ توافقٍ، ولو استَعضنا عن المدخلات الأخلاقية بمدخلاتٍ حقوقيةٍ، فلا توجد ضمانة بأن هذا سيكون مقبولًا عبر الثقافات. وهذا يثير أسئلةً مُلِحَّةً حول الأطر الأخلاقيَّة التي ستحكم أنظمة الذكاء الاصطناعي؟ ما هي القِيَم والأنظمة الأخلاقية التي ستُوظَّف لتوجيه تصرفات هذه الآلات الذكية؟ وكيف سنضمن أن القرارات التي تتَّخذها هذه الأنظمة تعكس اعتبارًا متوازنًا للمنظورات الأخلاقيَّة المتنوعة الموجودة

في مجتمعنا العالميّ؟ هل علينا أن نحل مشكلة الأخلاق في الذكاء البشري قبل أن نفكر في حلول لها في الذكاء الاصطناعي؟ تحتاج الإجابة عن هذه التساؤلات إلى حواراتٍ موسَّعةٍ وانفتاحٍ قد لا يكون متوفِّرًا الآن، ولكنه ممكنٌ إذا توفَّرت الإرادةُ.

وعلى صعيد آخر، تلقي قضية "الملاءمة" بظلالها على جانبٍ آخر من جوانب المناقشة حول الأخلاق والذكاء الاصطناعي، وخاصةً عندما نفكر في مستقبل أنظمة الذكاء الاصطناعيّ الفائقة الذكاء، التي قد تعمل بشكلٍ مستقلٍّ، وتنطوي على عملياتٍ دون إشرافٍ بشريٍّ. وهذا يثير تساؤلاتٍ مهمةً حول الأُطُر الأخلاقيَّة التي ينبغي أن تحكم مثلَ هذه الأنظمة. ويزعم بعض الخبراء أن المبادئ الأخلاقية التي توجِّه السلوك البشري –المتجذرة في التعاطف، والمعايير الاجتماعية، والفلسفة الأخلاقية– قد لا تكون كافيةً أو مناسبةً لأنظمة الذكاء الاصطناعيّ التي تعمل بشكلٍ مستقلٍّ. ويكمن التحدِّي في تحديد ما يشكِّل "الأخلاق المناسبة" لهذه الأنظمة المتقدِّمة؛ لأن المشهد مُعقَّدٌ وغير واضحِ المعالم. غالبا ما تشير المناقشات إلى سيناريوهاتٍ واقعيةٍ تسلِّط الضوء على المعضلات الأخلاقيَّة التي تواجهها الأنظمة المستقلَّة. وثمة مثالٌ على ذلك يكثر تداوله وهو يتعلق بالخيارات التي يتعين على المركبات ذاتية القيادة اتخاذها في حالات الحوادث الحرجة. فعندما تواجه تصادمًا مُحتَمَلًا، هل ينبغي للمركبة أن تعطي الأولويَّة لسلامة ركَّابها من خلال الانحراف لضرب المشاة، أم ينبغي لها أن تصطدم بجدار، وبالتالي تخاطر بحياة ركَّابها؟ إن هذا المأزق الأخلاقيَّ يوضِّح التوتُّر بين الأخلاق النفعيَّة من ناحيةٍ، وتعظيم السلامة العامة والقواعد الأخلاقية التي تثمن واجب حماية الحياة البشرية من ناحيةٍ أخرى. وفي حين قد يتصارع السائق البشري مع مِثل هذا القرار، فإنه غير مُقيَّد باستجاباتٍ مبرمجةٍ مسبقًا، بل يمكنه أن يَزِن الموقف في الوقت الحقيقي، مستعينًا ببوصلته الأخلاقيَّة واستجاباته العاطفيَّة. ومن الأمثلة الأخرى الشائع تداولها يتعلق بنشر الروبوتات القاتلة في الحروب أو عمليات مكافحة الإرهاب. وقد صُمِّمت هذه الآلات لغرض صريح هو القتل، وهو ما يثير تساؤلات أخلاقية عميقة حول طبيعة أفعالها والمساءلة عن تلك الأفعال. ويثير احتمال اتخاذ الآلات لقراراتٍ تتعلق بالحياة أو الموت دون تدخُّلٍ بشريٍّ مناقشاتٍ حاسمةً داخل المجتمع حول الآثار الأخلاقية لمثل هذه التقنيات. وإذا قبلنا أن أنظمة الذكاء الاصطناعي قد تكتسب الاستقلال في اتخاذ القرار وإنشاء السيناريوهات، فهل يعني هذا أنها سوف تمتلك أيضًا شكلًا من أشكال الوكالة؟

وفي هذا السياق، تشير الوكالة الأخلاقية إلى القدرة على التصرُّف بشكلٍ مستقلٍّ، واتِّخاذ الخيارات على أساس مجموعةٍ من القِيَم أو المبادئ، أي أن "يُوكَل" إلي أنظمة الذكاء الاصطناعي وضع قواعد أخلاقية ذاتية تلائمها. إذا تم تصميم أنظمة الذكاء الاصطناعي للعمل بشكلٍ مستقلٍّ، فإن السؤال الذي يطرح نفسه هو: هل يجب أن

تكون مُشبعةً باعتباراتٍ أخلاقيةٍ مماثلةٍ لتلك التي توجِّه السلوكَ البشريَّ؟ وعلاوةً على ذلك، إذا كانت هذه الأنظمة قادرةً على اتِّخاذ قراراتٍ لها آثارٌ أخلاقيةٌ كبيرةٌ، فمن المسؤول عن أفعالها؟ المطوِّرون، المستخدِمون، أم الذكاء الاصطناعي نفسه؟ وبينما نتنقل عبر هذه المشاهد الأخلاقية المعقَّدة، يصبح من الواضح بشكلٍ متزايدٍ أن تطوير أنظمة الذكاء الاصطناعي الفائقة الذكاء يتطلب إعادة تقييم أُطُرِنا الأخلاقية. يجب أن نأخذ في الاعتبار ليس فقط العواقب المباشِرة لهذه الأنظمة، ولكن أيضًا العواقب المُحتَمَلة التي قد تترتَّب عليها. إن الأخلاق في هذه الحالة لن تكون مُدخَلات من عالم البشر، وإنما نتاج عملياتٍ تظهر في صورة مُخرَجاتٍ. وحتى في حالة تدخُّل البشر لتغذية أنظمة الذكاء الاصطناعي بأخلاقياتٍ معيَّنةٍ فما هي طبيعة اختياراتنا؟ ومن يقرِّرها؟ ولمصلحة من؟

ويبقى السؤال: ما هو موقفنا حيال هذا المسار العالميّ؟ وعندما أقول "موقفنا" أدرك أن هذه الكلمة تحتاج إلى مزيدٍ من التحديد؛ فثمة مواقف متعدِّدة بحكم المصلحة، والتوجُّهات الأيديولوجية والسياسية والأخلاقية، ولذلك فأنا أتحدث عن التيارات الإنسانية، التي وإن لم تتَّفق على مفهومٍ واحدٍ للأخلاق، فعلى الأقل تَعِي أن هناك مصلحةً مشتركةً في مواجهة وتدارك مخاطر أنظمة الذكاء الاصطناعي. ولحسن الحظ أن أنصار هذا التيار لهم حضور في الأوساط البحثية والأكاديمية والمجتمعات المدنية والهيئات الدولية وفى المؤسسات الدينية... وغيرها. ومع ذلك، فإن المقصود هو هذه الأطراف في سياق المنطقة العربية بالأساس. ومن الواضح أن تحدياتنا تنبع في المقام الأول من أنفسنا قبل أن تكون تحدياتٍ خارجيةً مرتبطةً بالتقدم الحاصل في تكنولوجيا الذكاء الاصطناعي. نحن نواجه قضايا تتطلب منا التأمل وتحمل المسؤولية التاريخية بشأنها، ومن ذلك:

أولًا: أن وضعنا الهامشي في المجالات العلمية والتكنولوجية يجعلنا عرضةً لتهديداتٍ سياسيةٍ واجتماعيةٍ ووجوديةٍ. فأبسط تطبيقات الذكاء الاصطناعي يمكن أن تشكِّل تهديدًا في الأوضاع الهشَّة. ففي حين تعرب الدول المتقدمة عن مخاوفها بشأن الذكاء الاصطناعي المتقدم والأيديولوجيات ما بعد الإنسانية المصاحبة له، فإن المجتمعات المُهمَّشة تكتفي، وأحيانًا تتفاخر، بموقعها الاستهلاكي لمنتجات الذكاء الاصطناعي الأساسية، وغالبًا ما تقع فريسةً لمنتجات دون المستوى، أو تلك التي تخدم في النهاية مصالح ضيقةً تتلاقى ومصالح محتكري التكنولوجيا والذكاء الاصطناعي.

ثانيًا: على الرغم من ادِّعاءاتنا بالتفوق الأخلاقي، فإننا ندرك تمام الإدراك أن معاييرنا الأخلاقية غالبًا ما تعمل كواجهاتٍ تخفي نقاط ضعفنا. هذه الواجهات، التي نتصورها عن طريق الخطأ على أنها معاقل ضد التهديدات

الخارجية، تعيق قدرتنا على التكيف مع متطلبات الحداثة والتقدم السريع في العلوم والتكنولوجيا. لذا فمن الضروري أن نعترف بعدم جدوى هذا النهج الثقافي في مواجهة الهجوم الشرس للتكنولوجيا والذكاء الاصطناعي. فهل نحن مستعدُّون لإعادة تقييم مواقفنا الأخلاقية بشكلٍ نقديّ ومسؤولٍ، وخاصةً فيما يتَّصِل بحقوق الإنسان، التي تمثِّل، بغضّ النظر عن تفضيلاتنا، إنسانيةً مُشتَرَكةً وملاذًا ضد العديد من المخاطر السياسية والبيئية والتكنولوجية؟

ثالثًا: في حين أننا قد نَستَغِل فوائد الذكاء الاصطناعي في مجالاتٍ مثل الرعاية الصحية وإدارة البيئة، فإن غياب حوكمةٍ رشيدةٍ بما تعنيه من مساءلةٍ ومشاركةٍ، قد يجعلنا عرضةً بشكلٍ خاصّ للجوانب الضارة للذكاء الاصطناعي، مُمثَّلة في كونها أدواتٍ للسيطرة والمراقبة وانتهاك الخصوصية. وإذا كانت المجتمعات الغربيَّة تمتلك الأُطُر التنظيميَّة اللازمة لإنفاذ القوانين والتأمل الجاد في حوكمة الذكاء الاصطناعي، ففي السياقات الخالية من مثل هذه الأُطُر، قد يعمل الذكاء الاصطناعي كآليةٍ لتقويض فرص الحوكمة الرشيدة.

وأختم كلامي بالتوصية الخاصة بـ"أخلاقيات الذكاء الاصطناعي" الصادرة عن المؤتمر العام لمنظمة الأمم المتحدة للتربية والعلم والثقافة (اليونسكو) في دورته الحادية والأربعين في الفترة الممتدة من 9 إلى 24 نوفمبر من عام 2021. فهذه التوصية يمكن أن تكون وثيقةً إرشاديةً للعمل المُشتَرَك حيث: ".. تبحث التوصية أخلاقيات الذكاء الاصطناعي بحثًا معياريًّا منهجيًّا يقوم على إطارٍ كليّ شاملٍ ومتعدِّد الثقافات ومتغيِّر القِيم ومبادئ وإجراءات مترابطة يمكن أن ترشد المجتمعات فيما يخص التصدي بطريقةٍ مسؤولةٍ للعواقب المعروفة وغير المعروفة لوسائل تكنولوجيا الذكاء الاصطناعي على البشر والمجتمعات البشرية، وكذلك على البيئة والنُظم الإيكولوجية؛ وترسي التوصية بالتالي الأسس اللازمة التي يستطيع البشر الاستناد إليها للحكم على وسائل تكنولوجيا الذكاء الاصطناعي وقبولها أو رفضها. وتعتبر التوصية الأخلاقيات أُسسًا مرنة للتقييم المعياري والتوجيه المعياري لوسائل تكنولوجيا الذكاء الاصطناعي، وتستند إلى وجوب صون كرامة الإنسان ووجوب ضمان السلامة والرفاهية ووجوب درء الضرر باعتبارها البوصلة التي يجب الاسترشاد بها، وباعتبارها مبادئ راسخة في أخلاقيات العلوم والتكنولوجي".